Labor_02

LINEAR REGRESSION WITH MULTIPLE VARIABLE

Hypothesis

$$\underset{m\times(n+1)}{X} = \begin{bmatrix} x_0^1 & x_1^1 & x_2^1 & \dots & x_n^1 \\ x_0^2 & x_1^2 & x_2^2 & \dots & x_n^2 \\ x_0^3 & x_1^3 & x_2^3 & \dots & x_n^3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_0^m & x_1^m & x_2^m & \dots & x_n^m \end{bmatrix}, \underset{(n+1)\times 1}{W} = \begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}, \underset{m\times 1}{Y} = \begin{bmatrix} y^1 \\ y^2 \\ y^3 \\ \vdots \\ y^m \end{bmatrix}$$

Hypothesis:

$$h_w(x) = w_0 + w_1x_1 + w_2x_2 + \ldots + w_nx_n$$

$$h_w(x) = w_0x_0 + w_1x_1 + w_2x_2 + \ldots + w_nx_n$$

Additional X_0=1 (BIAS)

X*W

Estimation

$$\underset{(n+1)\times 1}{W} = \begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$$

$$\underset{m\times(n+1)}{X} = \begin{bmatrix} x_0^1 & x_1^1 & x_2^1 & \dots & x_n^1 \\ x_0^2 & x_1^2 & x_2^2 & \dots & x_n^2 \\ x_0^3 & x_1^3 & x_2^3 & \dots & x_n^3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_0^m & x_1^m & x_2^m & \dots & x_n^m \end{bmatrix} \quad \begin{bmatrix} w_0x_0^1 + w_1x_1^1 + w_2x_2^1 + \dots + w_nx_n^1 \\ w_0x_0^2 + w_1x_1^2 + w_2x_2^2 + \dots + w_nx_n^2 \\ w_0x_0^3 + w_1x_1^3 + w_2x_2^3 + \dots + w_nx_n^3 \\ \vdots \\ w_0x_0^m + w_1x_1^m + w_2x_2^m + \dots + w_nx_n^m \end{bmatrix}$$

Cost and Grad

Cost function:

$$C(W) = C(w_0, w_1, \ldots, w_n) = \frac{1}{2m}\sum_{i=1}^{m}(h_w(x^i) - y^i)^2$$

Gradient Descent:

$$w_j := w_j - \mu\frac{1}{2m}\sum_{i=1}^{m}(h_w(x^i) - y^i) \cdot x_j^i$$

For j on features
j=0…n

Data preparation

Feature Scaling

Mean Normalization

=> Faster convergence

$$x = \frac{x - mean(x)}{std(x)}$$

HINT:

$$v = \begin{bmatrix} v_0 \\ v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

$$v' = \begin{bmatrix} v_0 & v_1 & v_2 & \dots & v_n \end{bmatrix} \quad \left[(v_0)^2 + (v_1)^2 + (v_2)^2 + \dots + (v_n)^2\right] \quad \rightarrow v'v = sum(v.\hat{}2)$$